Which Factors Affect Access Network Performance?

Srikanth Sundaresan[1], Nick Feamster[1], Lucas Dicioccio[2,3], Renata Teixeira[3]
[1]College of Computing, Georgia Tech, [2]Thomson Labs, [3]CNRS/UPMC Paris Universitas
{srikanth.sundaresan,feamster}@cc.gatech.edu, lucas.dicioccio@thomson.net, renata.teixeira@lip6.fr
Georgia Tech Technical Report GT-CS-10-04

ABSTRACT

This paper presents an analysis of the performance of residential access networks using over four months of round-trip, download, and upload measurements from more than 7,000 users across four ADSL and cable providers in France. Previous studies have characterized residential access network performance, but this paper presents the first study of how access network performance relates to other factors, such as choice of access provider, service-level agreement, and geographic location. We first explore the extent to which user performance matches the capacity advertised by an access provider, and whether the ability to achieve this capacity depends on the user's access network. We then analyze the extent to which various factors influence the performance that users experience. Finally, we explore how different groups of users experience simultaneous performance anomalies and analyze the common characteristics of users that share fate (e.g., whether users that experience simultaneous performance degradation share the same provider, city). Our analysis informs both users and designers of networked services who wish to improve the reliability and performance of access networks through multihoming and may also assist operators with troubleshooting network issues by narrowing down likely causes.

1. INTRODUCTION

Residential access networks, such as Digital Subscriber Line (DSL) and cable are seeing steady deployment. Over the past decade, Internet usage has grown by more than 3.5 times, to about 1.6 billion users, about 300 million of which are broadband subscribers [16]. Despite this increasing penetration and users' increasing reliance on broadband networks for critical and performance-sensitive applications (*e.g.*, voice, video, gaming), very little is known about the factors that affect the throughput and latency experienced by users of broadband access networks. A better understanding of how the choice of Internet Service Provider (ISP) and service-level agreement (SLA) affect performance can help users make better decisions to improve both reliability and performance. In addition to ISP and SLA, understanding how performance varies per city can also help the designers of networked services (*e.g.*, overlay networks, content distribution networks) to decide where to replicate content and services to avoid paths that experience simultaneous performance degradations.

In this paper, we study how factors such as geography (*i.e.*, city), service provider, and service-level agreement affect the network performance that residential network users experience. Previous studies have mainly focused on first-order characterization of access networks [9, 24]; we build on these studies by exploring the relationship between a wide range of factors and user performance metrics, including throughput and latency. Specifically, we study the following questions:

- **Does performance match SLA?** (Section 4) We study whether the performance that users observe in practice matches the capacity that an access provider (ISP) advertises. We look at various parameters that illustrate how often users can expect to receive promised performance.
- **Which factors affect user performance?** (Section 5) We study how factors such choice of ISP and SLA, the user's city and neighborhood, and local time of day affect the access network performance For example, we study the extent to which users in the same city but different ISPs experience similar performance.
- **How does performance correlate across time?** (Section 6) We explore the question of shared fate among users. For example, we study whether users in the same city but different ISPs experience anomalies in performance at the same time.

We study these questions using active measurements collected over 4.5 months from thousands of unique residential user machines in France. This dataset was collected by Grenouille [14], a nationwide project to measure the performance of access links in France. Grenouille members install a client that performs ICMP probes and FTP uploads and downloads to a fixed set of servers every 30 minutes. When users install the client, they also enter their ISP, SLA, and city, providing valuable metadata about the measurements themselves.

The Grenouille dataset is distinct from previous datasets in three interesting ways: (1) Its measurements are taken directly from residential users' networks, as opposed to from somewhere outside the home network. (2) All measurements contain detailed metadata about the performance measurements, which allow us to analyze them according to various factors. (3) The dataset is larger than many existing studies, but, more importantly, it is heterogeneous, representing a set of users with a wide range of ISPs, SLAs, and geography. These three unique factors allow us to perform what we believe is the first study into the relationships between these underlying factors and the performance that a user achieves on an access network.

Despite the wealth of this dataset, we faced several challenges in performing this analysis. First, the Grenouille client only probes the network when the user is online, so the data has sizeable gaps for each user. These interruptions in the data collection make it difficult to correlate observed performance across users and consequently assess the underlying causes of differences in performance. To solve this problem, we aggregate data across users in the same SLA, ISP, and city to get a near-continuous timeseries for each performance metric of interest. Second, although the widespread client deployment provides a large, diverse dataset, it also makes it difficult for us to modify the data that the clients are collecting. This characteristic limits the conclusions that we can draw from our study; for example, we were not able to perform fine-grained performance measurements. Third, some of the performance measurements (*e.g.*, round-trip times from the ICMP probes) are averaged across a number of probes, and only performed once every 30 minutes. The aggregated nature of the measurements can prevent us from detecting certain types of performance characteristics. Fi-

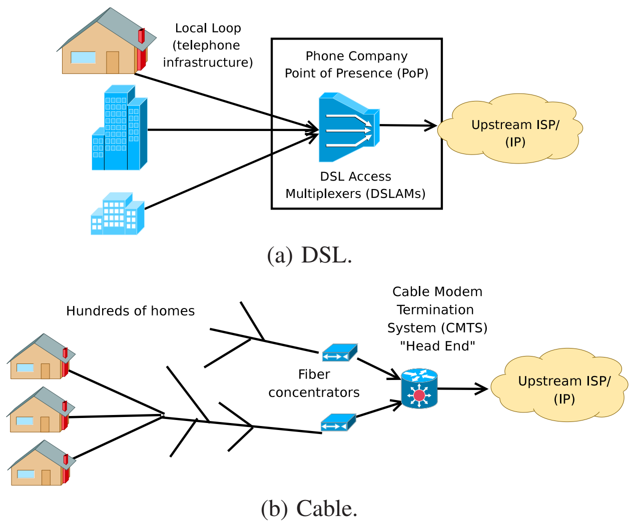


Figure 1: Access network architectures.

nally, the servers and files used for FTP downloads and uploads differ according to the measured ISP and SLA, so we must normalize such differences when comparing FTP measurements across ISPs.

Our analysis uncovers many interesting findings that shed light on access network performance. We observe that many users never achieve the SLAs advertised by their access ISPs, and that the for an ISP to meet its SLA depends on the advertised rate—lower rates are easier to deliver than higher rates. Moreover, we see that access network performance depends on SLA and ISP more than geography and time-of-day, but that the importance of ISP decreases for higher upload speeds. Finally, we observe that groups of users experienced *correlated* performance fluctuations, and that these correlations occur mostly for users who share the same access ISP, as opposed to other factors (e.g., city). We believe these findings both offer a deeper understanding into access network performance and open several avenues for future work in access network troubleshooting and performance.

2. RESIDENTIAL ACCESS NETWORKS

This section provides background on residential broadband access networks that we measure in this study. We provide a brief overview of Digital Subscriber Line (DSL) and cable networks and then explain how a user's choice of SLA and local configuration can potentially affect performance.

Cable and Digital Subscriber Line (DSL) networks are two common types of access technologies. DSL networks use telephone lines; subscribers have dedicated lines between their own DSL modems and the closest access multiplexer ("DSLAM"). The DSLAM multiplexes data between the access modems and upstream networks, as shown in Figure 1(a). The most common type of DSL access is asynchronous DSL (ADSL), which provides different upload and download rates. In cable access networks, groups of users send data over a shared medium (typically coaxial cable); at a regional *headend*, a cable modem termination system (CMTS) receives these signals and converts them to Ethernet, as shown in Figure 1(b). Cable operators often shape each user's traffic to enforce fairness. The physical connection between a customer's home and the DSLAM or the CMTS is often referred to as the *local loop*.

Users contract a service-level agreement (SLA) with their provider for some *maximum* capacity. The ITU-T standardization body establishes that the achievable rate for ADSL 1 [17] is at most 12 Mbps download and 1.8 Mbps upload. The ADSL2+ specification [18] extends the capacity of ADSL links to at most 24 Mbps download and 3.5 Mbps upload. Although the ADSL technology is theoretically able to reach these speeds, there are many factors that limit the capacity in practice. An ADSL modem negotiates with the DSLAM the rate to operate (often called the *sync rate*); this rate depends on the quality of the local loop, which is often determined by the user's distance to the DSLAM. The maximum IP link capacity is lower than the sync rate because of the overhead of protocols at lower layers. The best SLA that an ADSL provider advertises usually represents the IP rate that customers can achieve if they have good connections to the DSLAM. As we see in Section 3, a provider can also offer other SLAs with lower rates. These lower rates are usually implemented by limiting the customer's traffic at the DSLAM (e.g., using weighted-fair queuing). In a cable network, the actual rate that a user receives will vary with the network utilization of other users connecting to the same headend. In this paper, we study how the performance that a user receives corresponds to the SLA they have received from their ISP.

A user's local configuration of the DSL modem can affect the network performance that they experience. ADSL users configure their modems to work in either *fastpath* or *interleaved* mode. These modes affect only the local loop. In *fastpath* mode, the DSL modem transmits bits to the DSLAM in the same order that they were received, which minimizes latency but often limits the effectiveness of Forward Error Correction (FEC) in the event of bit errors. For this reason, ISPs only recommend that users configure fastpath if the line quality is good. *Interleaved* mode is more robust to errors because the modem interleaves bits of different packets (one cell will have the first bit of the first packet, first bit of the second, and so on; the second cell will have the second bit of the first packet, second bit of the second packet, etc.). In interleaved mode, the loss of one cell only affects a single bit of any packet, making it possible for FEC to recover from the error. Thus, interleaved mode offers better error correction properties at the cost of higher latency. In this paper, we quantify the extent to which a user's configuration of the DSL modem affects download performance and latency.

Another important distinction in DSL and cable networks is that the owner of the local loop might differ from the access network operator. Historically, each country had a local telecom operator (for instance, the "Baby Bells" in the US or France Telecom in France), which owns the local loop. Since the mid-1990s, governments in many developed nations have introduced a regulatory framework to provide "local-loop unbundling", which gives competing network operators access to the local loop. For example, in the United States, the Telecommunications Act of 1996 required the incumbent local exchange carriers (ILECs) to lease their equipment to competitive local exchange carriers (CLECs) [4]. In France, local-loop unbundling started in 2003 [22]. Orange,(previously called France Telecom) owns the local loop, but competing carriers can install their own DSLAMs in Orange's points of presence (PoPs). The unbundling process is often slow. Not all Orange's PoPs are open to competitors yet. In these areas, other service providers use Orange's DSLAMs to connect to their network; this configuration is called *bundled* service. When users connect directly to their provider DSLAM, this service is called *unbundled*. In this paper, we also compare the performance of a single ISP when it offers a bundled versus an unbundled service.

3. THE DATA

In this paper, we use data collected by the Grenouille project, which measures residential access network performance for ISPs across France. We first describe the process of collecting data within the Grenouille project and the resulting dataset. Then, we discuss the various challenges we faced in processing and analyzing the data, because we did not control its collection.

3.1 Data Collection

In 2000, a group of volunteers started the Grenouille project to monitor the performance of residential access providers in France. This project now has thousands of members across all major French cities and Internet service providers. To participate in this project, users download the Grenouille client, which runs periodic tests to measure their provider's performance. The Grenouille client runs on Windows, MacOS, and Linux. The client performs three types of periodic measurements: round-trip time (RTT), average FTP download rate, and average FTP upload rate. After collecting these statistics, the client sends these results to a central server. The server aggregates these measurements to construct statistics for each ISP, SLA, and city. Users can then view the performance statistics for each ISP at `www.grenouille.com`.

To configure the client, users create an account with Grenouille and enter information about their connection. Users enter the city where they live, their ISP, and their SLA. Based on these parameters, the Grenouille server configures the measurements that the client should perform. The destination for both the RTT probes and the FTP uploads is always a Grenouille server (except for Numericable, which has its own upload hosts). The source of the FTP download depends on the ISP. The client usually downloads the files from a server inside the user's own ISP. If the ISP does not have an FTP download server, then clients download the file from the Grenouille server. The size of the file varies according to the SLA so that the client does not congest the network for users with slower connections.

Each Grenouille client performs measurements using ping and FTP; the client first checks that the network card is idle before performing measurements This minimizes interference from any traffic or activity at the end host that might affect the measurements. Every 30 minutes, a client performs one FTP download and one upload to estimate download and upload speeds and estimates the RTT and loss rate with the average of ten consecutive pings. Clients periodically send the result of these measurements to the server along with time that elapsed between when the measurements were performed and the time when client sends the report to the server. The server then timestamps the measurement by subtracting this time difference from the time at which it receives the message. This mechanism allows the server to synchronize the measurements from the clients without requiring the clients' clocks to be synchronized. The server then truncates the resulting timestamps to the nearest 30-minute timestamp and averages the performance measurements it received over that period from the client. The server then stores these averages in a database.

3.2 Data

This paper analyzes 12 months of measurements from the Grenouille project, from January 1, 2009 to December 31, 2009. All information concerning user accounts have been removed from this dataset, but we do have a unique identifier for each user. In addition to the raw performance measurements, we also have metadata about each user; specifically, we have the user's upstream ISP, SLA, and city. This dataset contains 7149 unique members; this number is small compared to the total number of broadband access subscribers in France (13.67 million in 2007 [16]), but Grenouille has clients in all ISPs and in all big cities in France.

Table 1 shows the number of unique users per ISP and SLA. We only study (ISP, SLA) pairs with more than 200 unique users. Orange, Free, and Neuf are ADSL providers; Numericable is a cable provider. For Free and Neuf, the table indicates whether each SLA is already unbundled or still uses Orange's DSLAMs. The table also presents for each ISP and SLA the measurement setup, *i.e.*, the location of the server used for the FTP download and upload as well as the size of the downloaded/uploaded files. Free, Neuf, and Numericable deploy servers inside their networks for the Grenouille FTP download measurements; Numericable also deploys a server for the FTP uploads. The size of the files used for FTP transfers are adjusted according to the SLA to avoid overloading the user's link. These files are sufficiently large to avoid any bias that could be introduced by the TCP slow start mechanism. Grenouille clients always send the ping measurements to the same Grenouille server. All Grenouille servers are located in an AS that has direct peering with all of these ISPs at the same Internet exchange point, which is located in Paris.

The two ISPs that have the most number of Grenouille clients are Orange and Free (these are also the two biggest ISPs in France). Table 2 presents the number of unique users per city for these two ISPs. The Grenouille data has users in more cities. We selected these five cities because they had more than one hundred users each for both Free and Orange. When presenting results that aggregate all users in an ISP or SLA, we use data from users in all cities.

For the Free network, we have access to metadata that maps each IP address to its associated DSLAM [8]. In contrast to the other access networks, Free has a full IP network: Each DSLAM is allocated an IP prefix and customers have static IP addresses. Free announces the address range for each DSLAM. This additional data allows us to examine the relationship between the user's DSLAM and their observed performance.

3.3 Challenges and Limitations

The Grenouille data offers measurement perspectives from thousands of users; unfortunately, the size of the deployment also means that we cannot control or change the measurements that clients collect, and the size of the dataset also means that the data is often aggregated or otherwise limited. We briefly discuss how the nature of the dataset imposes several challenges and limitations.

First, the measurements from each Grenouille client are *not continuous*, because each user is not always connected, and users may uninstall or stop the client at any time. When we perform per-user analysis, we take these gaps into account. Fortunately, much of our analysis focuses on groups of users for a specific ISP, SLA, and city; grouping users into these bins allows us to overcome the gaps in the data that may exist from any individual client. Second, the measurements are *coarse*: Clients perform measurements every 30 minutes, timestamps are rounded to 30-minute bins, and the server aggregates ping measurements over ten consecutive trials, which prevents us from performing any fine-grained analysis. Third, the client uses a *limited set of measurement tools* that are already available on hosts: ping and FTP; this choice allows the client to achieve wide deployment but means that we do not have access to tools that could provide more precise measurements. Fourth, as previously mentioned, clients probe only a limited set of destinations, and these destinations may be slightly different for each ISP, which limits our ability to compare performance across ISPs.

4. DOES PERFORMANCE MATCH SLA?

In this section, we explore whether the network performance of residential customers matches the maximum advertised by ISPs. After describing our method to measure the network performance per customer (Section 4.1), we study how often the performance

ISP	SLA	Status	Download (kbps)	Upload (kbps)	Unique members	Download host	Upload host	Download size (kB)	Upload size (kB)
Orange	ADSL-512	N/A	512	128	252	grenouille	grenouille	1024	100
	ADSL-1024	N/A	1024	128	426	grenouille	grenouille	1024	100
	ADSL Max	N/A	8192	800	1834	grenouille	grenouille	8192	1024
	ADSL2+	N/A	18432	1024	908	grenouille	grenouille	8192	1024
Free	ADSL-2048	bundled	2048	128	308	Free	grenouille	2048	300
	ADSL-10M B	bundled	10270	1024	553	Free	grenouille	8192	1024
	ADSL-10M U	unbundled	10270	1024	2303	Free	grenouille	8192	1024
	ADSL2+	unbundled	28672	800	3284	Free	grenouille	8192	1024
Neuf	ADSL-2048	bundled	2048	256	215	9tel	grenouille	2048	250
	100% neufbox	unbundled	16998	1024	971	9tel	grenouille	8192	1024
	MaxiDSL	bundled	20480	800	411	grenouille	grenouille	8192	1024
Numericable	30 Mega	N/A	30720	1024	719	Numericable	Numericable	16384	1024
	30 Mega (ex Noosnet)	N/A	30720	1024	229	Numericable	Numericable	8192	1024
	100 Mega	N/A	102400	5120	754	Numericable	Numericable	16384	1024

Table 1: The ISPs and SLAs that we studied.

City	Free SLAs				Orange SLAs			
	ADSL-2048	ADSL-10M U	ADSL2+	ADSL-10M B	ADSL-1024	ADSL Max	ADSL-512	ADSL2+
Lyon	12	103	105	27	10	79	17	46
Toulouse	26	76	88	68	13	171	5	42
Paris	4	139	294	7	0	103	0	63
Rennes	8	35	29	41	12	48	9	28
Bordeaux	31	61	79	6	9	101	11	27

Table 2: Geographic spread of users in Free and Orange.

that a user perceives achieves the SLA (Section 4.2).

4.1 Method

To explore the extent to which the user's achieved performance deviates from the promised SLA, we must devise metrics that reflect the performance users experience vs. the performance that their ISPs promise them via SLAs. We reflect this achieved performance in terms of two ratios: (1) the ratio of the user's 95th-percentile performance to the advertised rate (SLA); (2) the ratio of the user's median download speed to the advertised rate (SLA). The first ratio helps us estimate the maximum achievable capacity for each SLA; the second helps us determine how close the "typical" user performance is to the advertised SLA.

95th Percentile Performance. Given FTP download and upload measurements over 12 months, we need to estimate the actual maximum capacity of each user. Intuitively, the maximum download and upload rate during this measurement period should indicate the download and upload capacity limit. The maximum rate may reflect outliers, so we consider the 95th-percentile of the distributions of FTP download and upload rates to capture the maximum capacity of a user. We denote this metric as *P95*; our analysis compares this value with the rate advertised for the SLA by using the *P95-advertised ratio*.

Median performance. We measure whether users typically achieve good performance by measuring the median download and upload rates for each user; we compare this value to the advertised SLA and call this metric the *median-advertised ratio*. This ratio is close to zero when a user rarely experiences the advertised performance and close to one when performance is often close to the advertised rate.

Consistency of performance. We also quantify the consistency with which each ISP was able to deliver the best performance a user can get. To do so, we measure the ratio of the median to the 95th percentile rates, for both upload and download. We call this ratio the *Median-P95* ratio.

Second, some users' measurements are sparse: specifically, some users sent less than a handful of reports to the Grenouille server during the period of our measurements. Computing the 95th percentile or the median of FTP download and upload for these users may bias our analysis, since these users do not have many measurements in the first place. Thus, we only consider users that have sent at least 100 reports during our measurement period. Applying these two filtering steps yields a total of xxx users across the four different ISPs and 14 unique SLAs.

4.2 Results

We now present the results of our analysis based on the metrics from the previous section. We study both the 95th percentile and median performance for multiple ISPs and SLAs. We explore whether specific ISPs meet their SLAs more often than others; we also explore these metrics for different SLAs in the same ISP and compare the metrics for bundled vs. unbundled SLAs.

Tables 3 and 4 summarize our results for each ISP-SLA pair. They present the average(standard deviation) of the P95-advertised, Median-advertised, and Median-P95 ratios for FTP downloads and uploads, respectively. These tables also show the median ratios per ISP (lines labeled "All"). Figure 2 offers a closer look into the differences between users of different ISPs. Additionally, Figure 2(a) presents the cumulative distribution function of the P95-advertised ratio of download speeds across users of each ISP; whereas Figure 2(b) shows the cumulative distribution of the median-advertised ratio of FTP downloads across users of each ISP. When these ratios are close to 1, the user's performance is closer to the advertised rate

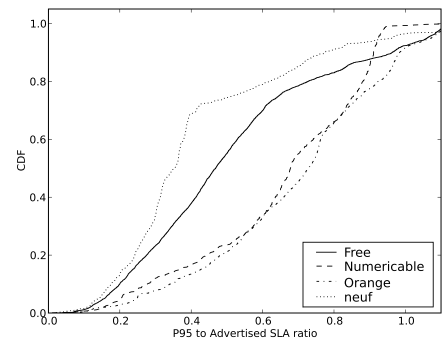
Result #1: Many users do not achieve their advertised SLA most of the time, particularly for download rates. Figure 2(a) shows that, for most users, the 95th-percentile of download speeds falls far short of the advertised SLA: for all ISPs, fewer than half of the users achieve 80% of what the SLA is advertising. Although it is expected that the download speeds will sometimes be lower than the advertised rate, these low values of P95 indicate that most users never get the advertised rate. The usual FTP download speeds is

ISP	SLA	P95 / Adv.	Median / Adv.	Median / P95
Orange	ADSL-512	0.95	0.80	0.89
	ADSL-1024	0.96	0.82	0.86
	ADSL Max	0.71	0.55	0.86
	ADSL2+	0.55	0.41	0.80
	All	*0.71*	*0.55*	*0.84*
Free	ADSL-2048	0.97	0.90	0.93
	ADSL-10M Bundled	0.76	0.69	0.91
	ADSL-10M Unbundled	0.74	0.53	0.81
	ADSL2+	0.45	0.34	0.81
	All	*0.55*	*0.41*	*0.82*
Neuf	ADSL-2048	1.13	0.89	0.90
	100% Neufbox	0.37	0.34	0.91
	MaxiDSL	0.28	0.22	0.88
	All	*0.35*	*0.29*	*0.89*
Numericable	30M	0.80	0.58	0.82
	30M (ex Noosnet)	0.66	0.54	0.79
	100M	0.56	0.38	0.78
	All	*0.67*	*0.50*	*0.80*

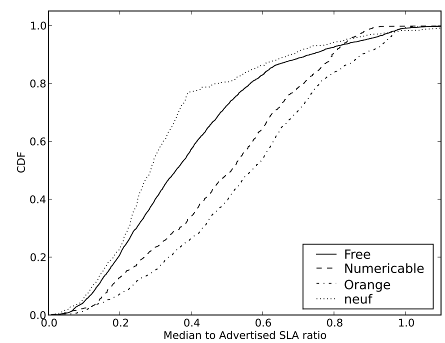
Table 3: FTP downloads: Median of ratios for all users in each ISP and SLA.

ISP	SLA	P95 / Adv.	Median / Adv.	Median / P95
Orange	ADSL-512	0.97	0.94	0.97
	ADSL-1024	1.04	0.91	0.86
	ADSL Max	0.94	0.86	0.97
	ADSL2+	0.79	0.75	0.98
	All	*0.88*	*0.80*	*0.97*
Free	ADSL-2048	1.06	1.03	0.99
	ADSL-10M Bundled	0.78	0.73	0.97
	ADSL-10M Unbundled	0.79	0.71	0.97
	ADSL2+	1.04	1.02	0.98
	All	*0.97*	*0.84*	*0.98*
Neuf	ADSL-2048	1.21	0.97	0.84
	100% Neufbox	0.64	0.62	0.96
	MaxiDSL	0.81	0.79	0.97
	All	*0.79*	*0.70*	*0.96*
Numericable	30M	0.94	0.92	0.99
	30M (ex Noosnet)	0.87	0.65	0.99
	100M	0.70	0.56	0.86
	All	*0.93*	*0.82*	*0.95*

Table 4: FTP uploads: Median of ratios for all users in each ISP and SLA.



(a) P95-Advertised Ratio



(b) Median-Advertised ratio

Figure 2: Download performance across different ISPs.

even lower than that. Figure 2(b) shows that more than half of the users of all the ISPs usually gets less than 60% of the advertised rate. Many factors affect the actual download rate of a user. In DSL, the DSLAM and the modem automatically negotiate a rate that depends on the quality of the local loop. Some DSLAMs may have many users, and consequently experience worse performance. In cable networks, the rate depends on the utilization of other users connected to the same headend.

In some cases, users achieve download rates that are *higher* than the advertised rate; we study these cases in detail. For ADSL SLAs that are less than or equal to 2048 kbps and for Numericable, the ISPs impose the limit using some shaping mechanism (*e.g.*, token bucket). Users experience bursts that are slightly above the limit. Neuf seems to set this limit slightly higher than 2048 kbps for ADSL-2048. The average download rate at any given time bin is often around 2400 kbps, but no user reports download speeds greater than 2500 kbps. For higher SLAs, ISPs usually advertise a limit that is slightly lower than the maximum achievable by the ADSL technology. Our results reflect this practice. The only exception is that some users of Free report download rates that are higher than the ADSL or ADSL2+ standard limits [17, 18]. Free does announce rates that are higher than the standard, but they announce an ATM rate. We are investigating if these few outlier values (11 users in total) represent some measurement artifact or some real technological difference in Free's DSLAMs and modems.

Result #2: Upload performance is more consistent than download performance. In contrast to download speeds, P95 of uploads are much closer to the advertised rates. Table 4 shows that for all ISPs the median user reaches upload speeds at least 80% of the advertised rate (*i.e.*, P95-advertised ratio larger than 0.8), and the median-advertised ratio is also higher than 0.8 for all ISPs except Neuf. Table 4 also shows that upload performance is much more consistent than download performance: the median-P95 ratio for upload performance is typically around 0.95 for most ISPs (suggesting upload speeds are consistent), whereas this value is never larger than 0.89 for any ISP for download performance.

The difference in upload and download performance results from the asymmetry of link capacity. As shown in Table 1, upload capacities are significantly lower than download: upload capacity varies from 128 kbps to 5 Mbps depending on the SLA, whereas download capacity varies from 512 kbps to 100 Mbps. Consequently, the bottleneck for uploads is most often the last hop capacity.

Result #3: An ISP's ability to consistently meet its SLA depends on the SLA. Figure 3 shows the median-advertised ratio for FTP download per user for each SLA of Orange. Interestingly, the distribution of this ratio is variable across different SLAs: specifically, SLAs with a lower advertised rate have a much higher median-advertised ratio than the SLAs with a higher advertised rate. We observe similar trends for all ISPs, as shown by the median values presented in Tables 3 and 4. To be clear, customers that subscribe to higher advertised rate still get better performance. For example, the

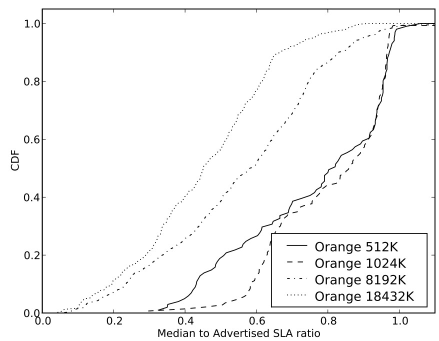


Figure 3: The median-advertised ratio for all SLAs from Orange.

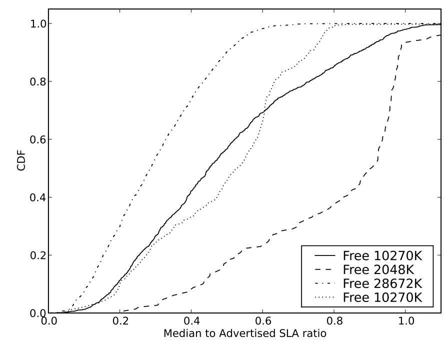


Figure 4: Med-advertised ratio for downloads for Free users.

median FTP download rate for Orange's ADSL-1024 is 932 kbps; this number is more than seven times higher for the ADSL2+ service.

The results in Figure 3 suggest that it is easier for ISPs to deliver lower rates. In the ADSL networks, this might be the case for a number of reasons. First, for the lower SLAs (which advertise download rates below 2 Mbps), the bottleneck both for downloads and for uploads is often the DSL link (both the DSLAMs and the backbone network are provisioned for the higher SLAs). Second, higher SLAs are limited by the capacity of the link. If the local loop has a high signal-to-noise ratio, the sync rate negotiated between the DSLAM and the modem will be lower than the advertised rate. For higher SLAs, downloads are also more likely to be limited by congestion at the DSLAM or the ISP backbone. Numericable, the only cable provider, advertises much higher download rates. The actual rates that users achieve are lower than the advertised rates not only because of congestion at the headend and backbone network, but also because of congestion in the local loop (given that the cable is shared).

Result #4: Unbundled service meets the advertised rate less often than bundled service. Figure 4 shows the median-advertised ratio for the same Free SLA for bundled vs. unbundled service. One would expect that when Free is operating its own DSLAMs (*i.e.*, in the unbundled service), it would deliver better performance. Our results, however, are not conclusive. Some users do achieve much better performance (sometimes even higher median FTP downloads and uploads than the advertised rates), but most users of unbundled service achieve the advertised rate less frequently than users of the bundled service.

In this section, we explored whether ISPs' advertised SLAs match the performance that they actually deliver and, to our surprise, found not only that performance often does not match the SLA, but also that access network performance is quite variable. This finding naturally leads us to our next question: what factors affect the performance that ISPs ultimately deliver to their end users? We explore this question in the next section.

5. WHICH FACTORS AFFECT PERFORMANCE?

In this section, we study the factors that affect the performance that users observe at their access links. We also quantify the extent to which these various factors affect performance. Specifically, we consider the following factors: *modem configuration* (*i.e.*, interleaving vs. fastpath); *geographical location* (city); *Internet Service Provider* (ISP); *Service Level Agreement* (SLA); the *DSLAM* for the user's access network; *time-of-day* and *day-of-week*. The unique characteristics of our dataset allow us to analyze the effects of these features on overall performance. Of course, there could be other features that also affect user performance (*e.g.*, a recent report suggested that even seasons and weather patterns might affect access link performance [21]). Our goal is not to provide an exhaustive list of features that affect performance, but rather to use the metadata that we do have to determine the effects of various factors on observed performance.

We perform this analysis with a sequence of tests. We first analyze the first-order properties of the features that affect performance (Section 5.1). Second, we perform classification and regression analysis using RuleFit [12], an ensemble learning algorithm that also provides insight into the relative importance of each of these features.

5.1 First-Order Properties

In this section, we examine the observable first-order effects of various features on performance. Interestingly, most of the features above have some observable effect.

Modem configuration: *Fastpath*, in contrast to the *interleaving* approach, is believed to improve RTT by approximately 20 ms. ISPs advise their users to enable this feature *only* if their local loop is high-quality, since low-quality links may actually experience worse performance with fastpath enabled (since many bits will be corrupted at the same time and error correction will not be able to recover).

Distinguishing users who have configured fastpath from those who have configured interleaving is difficult; the metadata we have for each user does not have any information about the configuration of the user's gateway. We only have the reported RTT to separate the two sets of users. We use the fact that there is a clear difference in RTTs achieved by fastpath users, as shown in Figure 5, which plots the measured round-trip times values of Free ADSL2+ users in Paris over two days. For each group of users, we compute a threshold RTT value that separates the two classes of users; in this case, that value is about 12 ms. We consider fastpath users as the group of users who report a ping value below the threshold but above outlier values [1]. Because users may change their configuration at any time, we also apply a moving window of one week to compute the set of users that meet the ping threshold.

[1]There were a handful of unusually low values, which we exclude from the analysis; we used a threshold of 2 ms.

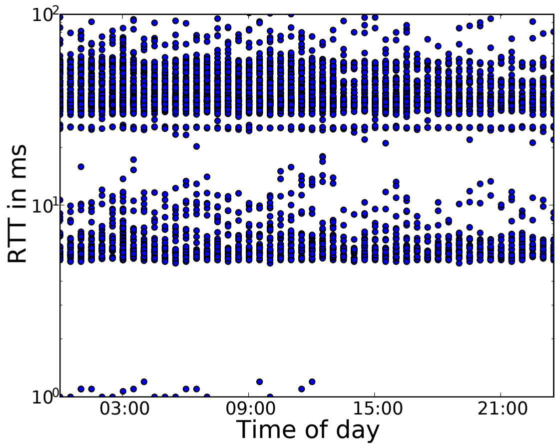


Figure 5: Scatter plot of RTT values reported by Free ADSL2+ users in Paris over two days aggregated over the time-of-day. The clear difference between users who have configured fastpath vs. interleaving enables use of threshold to separate Fastpath and Interleaved users.

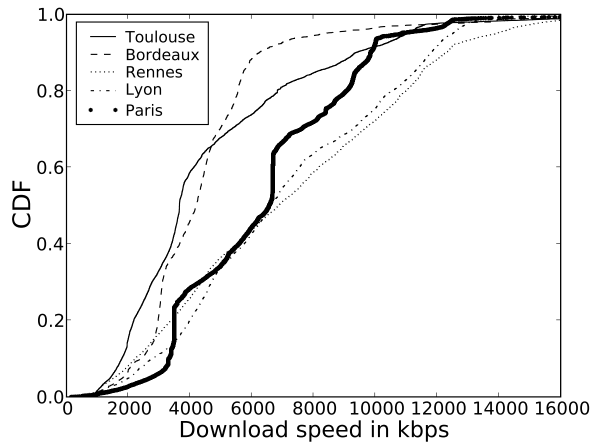


Figure 7: Effect of city: For the same ISP and SLA, user download performance can vary dramatically depending on city. This plot shows user download times for Orange ADSL2+ for different cities.

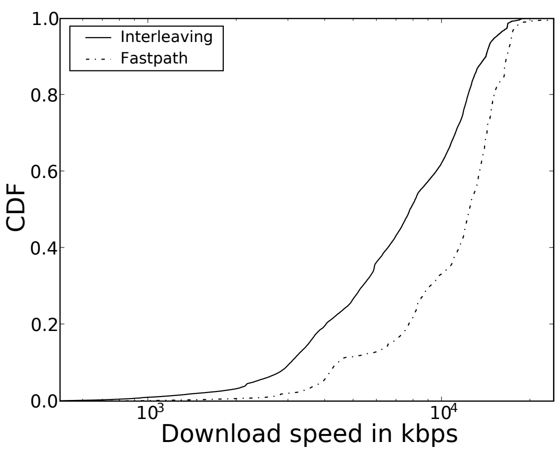


Figure 6: Effect of modem configuration: Users who have configured fastpath achieve better download speeds.

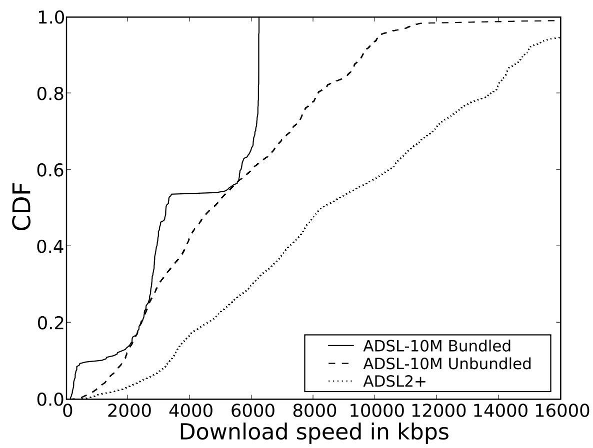


Figure 8: Effect of SLA: Users in a given city (Paris) within the same ISP (Free) see different performance.

Figure 6 shows the CDF of the download speeds obtained by fastpath and interleaving users. Fastpath users obtain higher speeds by a significant margin. There may be a bias here because in the general case, Fastpath is turned on only if the connection quality supports it.

City: To determine whether a user's city might affect performance, we plot the CDF of the download performance seen by users in the same ISP and SLA (Orange ADSL2+) across different cities. Figure 7 shows that the performance that a typical user for some ISP might experience across cities. We see that in some cases, these values vary quite significantly. For this SLA, users in Bordeaux and Toulouse experience significantly lower performance than users in Paris, Lyon or Rennes.

ISP: Figure 2 shows that the user's choice of ISP can affect whether the performance the user obtains matches the service-level agreement promised by the ISP. For example, the median R ratio was as low as 0.69 for some SLAs and has high as 0.97 for others. Here, we briefly explore some examples where users who are in the same city but use different ISPs might experience different performance. Figure 9 shows a comparison of FTP downloads for groups of users for different ISPs across Paris, for comparable SLAs. We saw similar variation in the upload rates.

Service Level Agreement: The SLA that the user buys should obviously affect the performance: Figure 8 shows how, for a single ISP and city, the choice of SLA can affect the user's observed download rates. This result is expected, although it is also interesting to note that even faster SLAs offer users slower performance at times, even when compared to the promised rates of slower SLAs.

DSLAM: We also examine the effect of DSLAMs on user performance. Figure 10 shows the CDF of the download speeds obtained by users connected to ten different DSLAMs. All the DSLAMs are in Paris, and all the users are subscribed to Free ADSL2+, with the Fastpath users filtered out. We see that performance varies quite

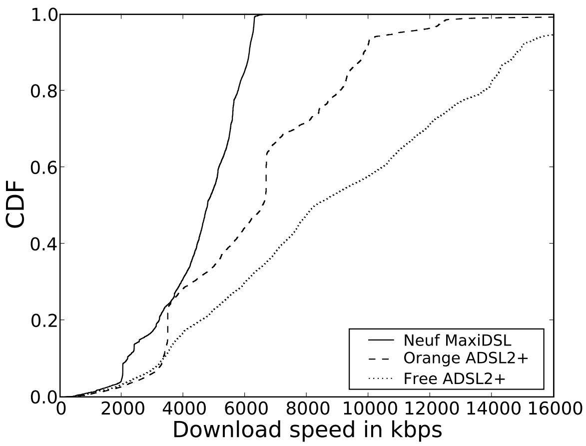


Figure 9: Effect of ISP: Users in a given city (Paris) can experience widely differing performance depending on their ISP, even with comparable SLAs.

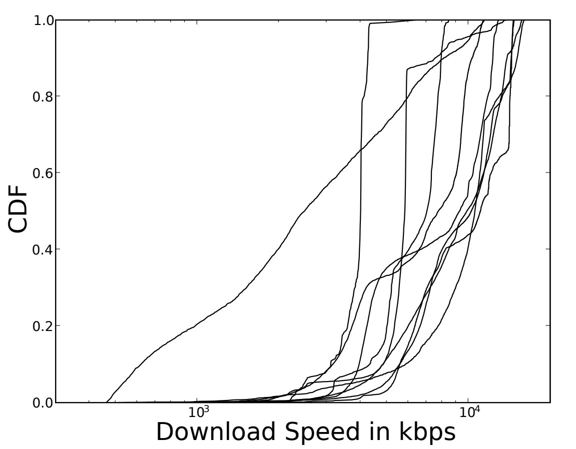


Figure 10: Effect of DSLAM: Download speeds for users connected to 10 different DSLAMs for the same ISP and city.

dramatically across different DSLAMs. One explanation for this is that the DSLAM that the user is connected to plays a major role in the performance she obtains. It might also be a function of the distance of the user from the DSLAM and the quality of the link.

Time of Day: We expect that performance will vary according to network utilization, and that overall network utilization will vary according to the local time of day. For example, Figure 11 shows that the median download speed obtained by Free ADSL2+ users varies according to time-of-day for users in Paris and Lyon, with performance peaking early in the morning and the lowest performance occurring during evening hours. However, for users in Toulouse and Bordeaux, the correlation is not so obvious.

Day of week: We expect that weekday traffic patterns will differ from weekend traffic patterns and, as such, a user with all other factors equal may experience different performance depending on whether it is the weekend or during the week. We did not observe any significance difference in the CDFs of download speeds across different days of the week.

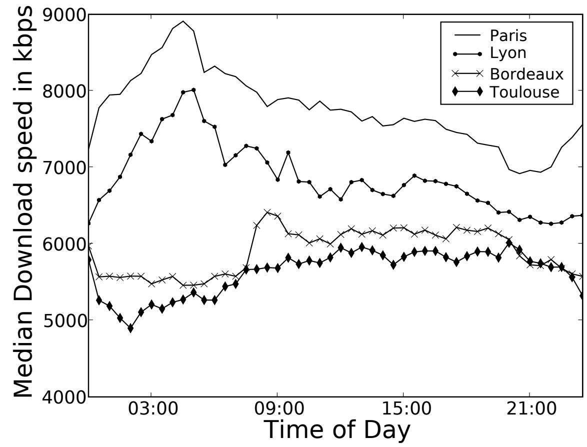


Figure 11: Effect of time of day: Download speeds across time-of-day for users in two different cities. Download speeds are generally higher in early-morning hours than they are later in the day.

5.2 Ranking of features: Ensemble learning

In this section, we explore the relative importance of each of the above features on user performance. In 5.1, we gained an intuition about the likely factors affecting performance; here we quantify the relative importance of each feature. We apply RuleFit, an ensemble learning technique, to better understand the relationship between each feature and the output variables.

Ensemble learning: RuleFit To gain more insight into the relationships between the input variables from Section 5.1, including information about the relative importance of each feature, we applied ensemble learning. Learning ensembles have emerged as one of the more popular predictive learning methods over the past decade. It combines simple functions of the input data ("base learners", which are indicator functions of input variables) into a decision tree. Each output variable (tree node) is defined by a "rule" that is the conjunction of the indicator variables from the root of the tree to that node. To improve the accuracy, the features themselves are also included as basis functions. Fast algorithms for minimizing the loss function [11] help make ensemble learning efficient. Ensemble learning models take the following form:

$$F(\mathbf{x}) = a_0 + \sum_{m=1}^{M} a_m f_m(\mathbf{x}) \qquad (1)$$

Where $\mathbf{x}$ are input variables derived form the training data (spatio-temporal features); $f_m(\mathbf{x})$ are different functions called ensemble members ("base learner") and M is the size of the ensemble; and $F(\mathbf{x})$ is the predictive output (in the case of regression, a numerical prediction of download or upload speed), which is based on a linear combination of ensemble members. Given the base learners, the technique determines the parameters for the learners by regularized linear regression with a "lasso" penalty to penalize large coefficients a_m. These coefficients help establish the importance of each variable. *RuleFit* [12] is a supervised *ensemble learning* classification method that produces a relative ranking of their importance to the output. Input variables that frequently appear in important rules or basic functions are deemed more relevant. The *RuleFit* paper explains the method in more detail [12].

We chose to apply RuleFit for two reasons: (1) it supports both classification and regression analysis on the input variables; and (2) it provides with a ranking of the *relative importance* of the input variables. RuleFit assigns variable importance based on the frequency of each feature's appearance in the important, more relevant predictors.

Input features We use all of the features from Section 5.1, as well as the round-trip time as input to RuleFit. We chose to include the round-trip time as an additional input feature for several reasons. First, all of the features from Section 5.1 are categorical; that is, they do not have numerical values. Providing one input variable to the regression that includes a numerical value helps improve prediction accuracy for the numerical output values (*i.e.*, FTP download and upload speeds). We have multiple observations for each combination of the input variables, including RTT, that we use as input for training. Because we also have a RTT value associated with most FTP measurements, we can use it as an input variable. In most cases, the server that the users ping from the FTP servers, but it turns out that RTT still is a useful indicator of performance. For the ISP Free, we also know the DSLAM that users connect to, but this information is not available for other ISPs; therefor we run the feature ranking across ISPs without the DSLAM information, and then run it for only Free users with the DSLAM information.

We run RuleFit using both classification and regression. For regression, the output variable is the actual reported upload/download value. Regression is straightforward, since the output variables that we are trying to predict (*i.e.*, download and upload speeds) are continuous. Because the classification and regression results are similar, both with respect to ranking and with respect to overall prediction accuracy, we show only the results from regression.

For clarity, and to simplify computation, we perform this analysis on 5 cities, 4 ISPs, 8 SLAs, and 3 times-of-day. We perform two types of regression analysis: regression over the entire dataset, and regression over specific ranges of the output variable. First, we run RuleFit on the *entire* set of filtered data, exactly as explained above. This analysis provides the overall relative importance of each input variable. Second, because we suspected that the relative importance of each feature might differ for different ranges of the output variables, we divide the data along the median and run RuleFit on each half of the data. We continue this process for four iterations, allowing us to perform RuleFit at various granularities and for 16 ranges of each output variable. This analysis allows us to determine how the relative importance of each feature varies for different ranges of performance. For example, there are if there are n total performance measurements and k bins, each bin contains n/k points, and the median value for the ith bin represents the $i \cdot 100/k$-th percentile. In this paper, we show the analysis for 16 bins, so the first bin includes $n/16$ data points whose median is about the 6th percentile of performance, and the sixteenth bin is about the 94th percentile. Performing RuleFit on a a subset set of data centered around such a lower or higher performance range allows us to assess whether a feature like city matters more or less for a certain range of values. The root-mean-square error of the prediction was generally an order of magnitude less than the median of the bin, but in some cases went up to about 50% of the median. This implies that the prediction accuracy is generally good, and the model built by Rulefit is quite good, too.

Result #1: RTT is most important; ISP, SLA, and city are also important. Table 5 shows the results from applying RuleFit's regression algorithms over the entire dataset without considering the user's DSLAM as a feature. As expected, the most important feature for both upload and download is RTT; the relative importance of RTT is expected because TCP throughput depends directly on RTT, and RTT is particularly important for short flows. The next most important feature for upload speed is ISP; for download speed, the next most important factor is SLA. The lower importance of SLA in upload speeds might be explained in part by the relative lack of variation in the upload speeds provided by ISPs among the higher end plans which dominates our dataset (due to the larger user base). Geographical location (City) is also an important predictor for both upload and download performance. This may be because performance may be constrained by infrastructural limitations like the distance from the nearest major PoP.

Rank	Upload	Download
1	RTT (100%)	RTT (100%)
2	ISP (51.3%)	Advertised rate (SLA) (85.5%)
3	City (35.2%	City (33.4%)
4	Advertised Rate (SLA) (11.4%)	Time of Day (8.4%)
5	Day of Week (0.2%)	ISP (7.4%)
5	Fastpath (0%)	Fastpath (2.7%)
6	Time of Day (0%)	Day of Week(0.2%)

Table 5: Ranking of features according to importance for the entire dataset.

Rank	Upload	Download
1	RTT (100%)	RTT (100%)
2	City (28.5%)	DSLAM (36.6%)
3	DSLAM (25.2%)	Advertised Rate (SLA) (33.7%)
4	Advertised Rate (SLA)(18.3%)	City (32.5%)
5	Fastpath (9.7%)	Time of Day (2.4%)
5	TIme of Day (2.1%)	Fastpath (0.4%)
6	Day of Week (0%)	Day of Week (0%)

Table 6: Ranking of features according to importance for the ISP Free, which also takes into account DSLAM information.

Result #2: A user's DSLAM is one of the most useful predictors of performance. The relative ranking of features with the DSLAM information included is shown in Table 6. Since we only have DSLAM information for Free users, the ISP factor becomes redundant, and so we remove it while running the tests. Here too, the most important feature for both upload and download performance is RTT (almost consistently 100%). The DSLAM is a dominant factor in both upload and download prediction, as important as the City and SLA.

Interestingly, the use of fast path does not affect the user's access network performance as much as other features. This result may simply result from the fact that we are including RTT as a separate feature, and this feature may subsume the feature of whether the user has enabled fast path. RTT values can still vary significantly, independently of whether a user has enabled fast path; the user's city and SLA may ultimately have a much greater predictive effect on download performance. We are exploring how to perform this regression with purely categorical variables, so that we can explore the relative importance of fastpath when RTT is removed.

Result #3: Temporal features are less important. Somewhat surprisingly, we found that temporal features such as time-of-day and day-of-week were less important features for predicting output, although time-of-day was more important for predicting download than it was for upload. This finding makes sense, since download performance may be more often affected by periods of high utilization (*i.e.*, lots of active users all downloading content) than upload performance (where it is less likely that all users will saturate the upload capacity at once).

Result #4: Relative importance for predicting upload speeds varies slightly according to the range of upload speed. Figure 12

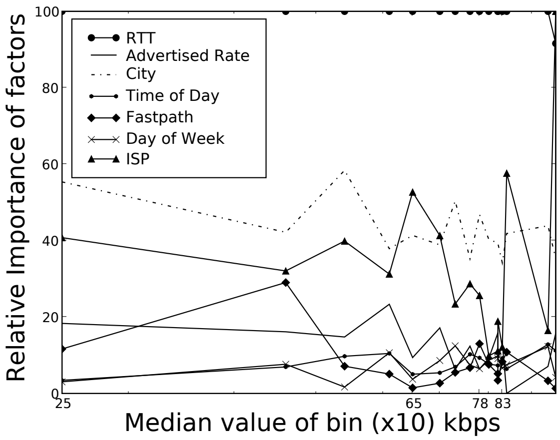


Figure 12: RuleFit ranking of features affecting *upload* performance (regression learning).

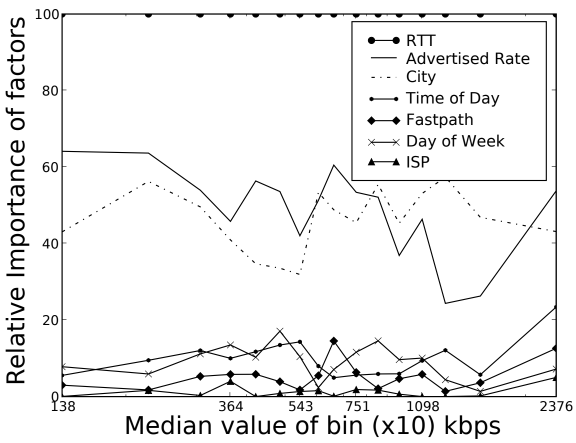


Figure 13: RuleFit ranking of features affecting *download* performance (regression learning).

shows the relative importance of the features when dividing the upload performance data into 16 continuous ranges. The left side of the plot shows the slowest download ranges; the rightmost side of the x-axis shows the highest values. For the lowest values of upload performance, ISP and city are about equally useful predictors. As upload performance increases, the importance of the city as a predictor stays roughly the same, while the importance of the ISP as a predictor of performance drops. For large upload speeds, the choice of ISP again becomes more important. Figure 13 shows the same type of ranking trend, but for *download* speeds. Once again, the two dominant features other than RTT are city and SLA. Surprisingly, we see for both download and upload, the time-of-day and day-of-week features are relatively insignificant. Although these relative importance rankings are not linear, they do suggest that the choice of ISP and SLA are consistently useful predictors for the upload and download performance that a user sees.

This section has offered some intuition regarding the importance of various features for predicting upload and download performance; in particular, we have seen that ISP and SLA are relatively more powerful predictors of performance than other features like geography and time-of-day. While this result would suggest that a user's choice of ISP and SLA is important in ultimately determining performance, we would also like to help users determine the factors that affect simultaneous performance fluctuations. In other words, we want to determine whether a group of users that experience simultaneous performance changes have any common features. By identifying whether groups of users that experience simultaneous performance fluctuations have any features in common, we might ultimately be able design methods to help users mask fluctuations and protect against failure.

6. HOW DOES PERFORMANCE CORRELATE ACROSS TIME?

In this section, we explore whether users of access networks experience performance fluctuations at the same time, and whether any certain underlying factors are responsible for this correlation. Determining whether user performance correlates according to geography or ISP, for example, can help users determine how to achieve better reliability in the face of failures. We explore this question by analyzing timeseries of performance measurements for users across multiple cities, SLAs, and ISPs. We first define a pairwise distance metric between timeseries; then we perform pairwise single-linkage clustering to identify groups of users that exhibit temporal correlation. Section 6.1 explains the data and our clustering-based analysis. Section 6.2 discusses our findings.

6.1 Method

We aim to analyze timeseries measurements across cities, SLAs, and ISPs to understand which groups of users experience performance fluctuations at the same time (in a sense, whether they "share fate"). For example, if an Orange user in Paris experiences a performance degradation, we would like to determine whether the performance degradation has more to do with the fact that the user is an Orange user or that the user lives in Paris. Determining whether performance fluctuations correlate more by ISP or by city provides useful information that help improve reliability: if users across the same ISP experience performance problems simultaneously, for example, then a user could improve performance by multihoming to multiple access ISPs. If, on the other hand, users across the same city experience correlated degradations, then improving reliability is more difficult, but could be achieved with an overlay network (e.g., by re-routing traffic through a different city).

To perform this analysis, we must overcome two significant challenges. First, we do not have a complete time series from a single user. This is because users don't usually keep their home machines running all the time. To solve this problem, we aggregate users that have the same city and SLA (and, hence, ISP) and create a timeseries by creating a data point that reflects the median performance value reported for each group of users for that time. Even aggregating users do not give us complete timeseries; we choose only those city-SLA pairs that have data points for at least 50% of the time period (April–June 2009). Second, because each report only has a granularity of 30 minutes, the timeseries are somewhat coarse-grained, which can make establishing correlation that occurs on small time-scales more difficult. However, we are able to perform longer term correlation with our timeseries.

Once we construct each timeseries, we infer groups of users that experience similar performance trends by defining a pairwise correlation coefficient for each pair of timeseries. Intuitively, the correlation coefficient gives us the strength of the relationship between the two series. Formally, the correlation coefficient is defined as the normalized cross-correlation at zero time lag; the normalized

Member 1	Member 2	Correlation coefficient
Lyon, Free ADSL2+	Paris, Free ADSL2+	0.76
Toulouse, Free 10M Unbundled	Paris, Free ADSL2+	0.56
Lyon, Free ADSL2+	Bordeaux, Free ADSL2+	0.56
Bordeaux, Free ADSL2+	Paris, Free ADSL2+	0.51
Lyon, Free ADSL2+	Toulouse, Free 10M Unbundled	0.50
Paris Orange ADSL2+	Paris Orange ADSL Max	0.47
Lyon Orange ADSL Max	Paris Orange ADSL Max	0.46
Lyon Orange ADSL Max	Paris Orange ADSL2+	0.42
Bordeaux, Free ADSL2+	Toulouse, Free 10M Unbundled	0.42
Lyon Orange ADSL2+	Paris Orange ADSL Max	0.40

Table 7: Correlation coefficients of the 10 most correlated download measurement timeseries.

cross-correlation function is:

$$R_{xy}(\tau) = \frac{E[(x(t) - \mu_x)(y(t-\tau) - \mu_y)]}{\sigma_x \sigma_y}$$

We use a pairwise correlation coefficient for each pair of timeseries as the *distance* between each group of users. The complement of the coefficient serves as a nearness metric, which we use for clustering

To cluster groups of related users, we use *single-linkage clustering*, a greedy hierarchical clustering algorithm that works by grouping the two closest leaf nodes, which are initially singleton clusters. At each step, a single leaf is added to the cluster, and this process iterates till all the elements are clustered under the *root*. The most important clusterings take place first, so following the progress of the clustering allows us to see which pairs are closest to each other.

6.2 Results

We construct timeseries using data from users in five SLAs and 4 cities (20 initial timeseries); we selected these SLAs and cities because their timeseries had more than 60% reports. For each group, we get the timeseries for RTT, upload, and download measurements. We construct timeseries based on three different samples: the 10^{th} percentile value, the median value, and the 90^{th} percentile value. We then apply the clustering to each set of readings (RTT, upload, download) separately.

The correlation coefficients show the strength of the relationships between the time series; it is a real number in $[-1, +1]$. A value of $+1$ implies a strong positive correlation, a value of -1 implies a strong negative correlation, and a 0 value implies independence. The results indicate strong correlation among the most correlated pairs of the download measurement timeseries, and weaker correlation between the RTT and the upload timeseries. The coefficients for the median timeseries are shown in Table 7; the coefficient for Lyon and Paris Free ADSL2+ for download measurements is 0.76, which is quite high. By contrast, the highest coefficient for the RTT and the upload measurements were 0.29 and 0.28, respectively.

The correlation coefficients show strong correlation among users, and the clustering results confirm it. with two distinct clusters forming: one for Free with 5 elements, with users in all four cities and across ADSL2+ and 10M Unbundled SLAs, and one for Orange with six elements, with all four cities and the ADSL2+ and ADSL Max SLAs. The results for upload and RTTs were not as definitive but still followed a similar pattern. We also noticed clustering among Neuf in the RTT and upload. These cities are not geographically close to each other. Unsurprisingly, we see some clustering of users in a single ISP in the same cities across *different SLAs*, which tends to suggest that users of different SLAs may experience the same failure modes. This finding makes sense, since congestion at a DSLAM or the failure of equipment might affect users across different SLAs similarly.

7. RELATED WORK

Our work studies the performance of residential access networks and how various factors affect this performance. The work we describe below has focused on *characterization*. Unlike our work, they have not studied either whether the achieved performance actually matches the promised SLA or the effects of various factors on access network performance.

Traffic measurements of ISP networks. Previous work has focused on measuring the performance of backbone ISPs using active measurement (*e.g.*, Keynote [20] and NetDiff [23]). Other work uses passive traffic measurements from a provider networks in Japan [6], France [26], and a large European DSL provider [24]. Although most of their characterization focuses on traffic patterns and application usage, these studies also infer RTT and throughput of residential users from traffic traces, but it is impractical to verify whether the performance users receive match the SLA with passive measurements alone. In fact, Siekkinen et al. [26] show that most traffic is limited by the application, not the network.

Measurement of access networks from remote servers. Other studies have characterized user performance with access only to a server connected to the Internet, probing user access links to infer the performance of these links [7, 9]. Active probing from a fixed set of servers can characterize a large number of access links, because each link can be measured from the same centralized server. Unfortunately, this method can produce inaccurate results because the measurement vantage point is not located on the access network itself. Additionally, none of these studies has analyzed the relative importance of various factors on user performance.

Measurement of access networks users' home networks. Grenouille provides the dataset used in this paper, places monitoring agents directly at the end-user's machine. Simpson and Riley [19] also advocate this approach, but they do not have as many users as Grenouille. The advantage of installing software at the end-host is that it measures the access network with the user's perspective. Han *et al.* [15] proposed an alternative approach to perform measurements from "inside" the home: by searching for open wireless access points, associating with these access points, and performing measurements from the researcher's laptop. Although this approach has the vantage point of inside a home without explicitly asking for end-users to install measurement agents, it does not scale. None of these projects have as many participants as grenouille. The dataset we use this paper is unique: it is the largest collected from users' homes and it contains extra information (such as the SLA and city) that are entered by users themselves.

Inference and troubleshooting tools. Other tools use a combination of passive and active measurements to help users troubleshoot performance problems. Netalyzr [25] measures the user's performance of commonly used protocols such using performance measurements from the client's browser. Network Diagnostic Tool (NDT) [5] and Network Path and Application Diagnostics (NPAD) [10] send active probes to detect issues with client performance; however, these tools do not compare performance across ISPs, and they do not seek to identify or compare the underlying causes of the performance fluctuations seen by users.

A variety of recent active and passive monitoring tools aim to help users identify whether their access ISP is the cause of performance degradation. Network Access Neutrality Observatory (NANO) [27] compares passively collected performance data from across users and aggregates this data according to groups of users that have similar features (*e.g.*, operating system, geography, etc.) to attempt to determine whether the user's access ISP is responsible for performance degradation. Glasnost performs active measurements to determine whether the user's ISP is actively blocking BitTorrent traffic [13]; however, the tool does not examine general performance fluctuations does it compare user performance across various factors (*e.g.*, geography).

Performance and reliability measurements of access networks. Multihomed Overlay Networks (MONET) seeks to improve the reliability of access networks for Web traffic by sending a user's traffic through multiple access ISP networks [3]; the paper observed that adding a second DSL link to a commodity connection significantly improved the overall reliability of the upstream connectivity. This finding is consistent with our results from Section 6, which show that performance fluctuations do not correlate across different ISPs. Akella *et al.* compared the performance improvements that result from multihoming to those that resulted from using overlay links [1, 2]; this study found that multihoming can eliminate almost all failures experienced by a singly-homed access network, which is also consistent with our findings in Section 6 that show that performance fluctuations are independent across different ISPs.

8. SUMMARY AND FUTURE WORK

This paper has presented the first detailed study of the factors that affect access network performance, using data from four ISPs, 14 SLAs and five cities across France. We explored three main questions: (1) Does access network performance match the SLAs advertised by ISPs?; (2) Which factors affect access network performance?; (3) How does performance correlate across time? Our experiments have highlighted several interesting findings. We observed that users' usual download rate is often less than half of the rate advertised by their access ISPs. The ability of an ISP to meet the SLA depends mainly on the advertised rate. We also observed that access network performance depends on SLA and ISP more than geography and time-of-day, but that the importance of ISP decreases for higher upload speeds. Finally, we saw that groups of users experienced *correlated* performance fluctuations, and that these correlations occur mostly for users who share the same access ISP, as opposed to other factors (e.g., city).

The findings from our study present several opportunities for future work. First, our finding that users of the same ISP often experience correlated performance fluctuations suggests that fluctuations may sometimes be due to infrastructure or provisioning effects. In future work, we plan to gather more fine-grained performance measurements and determine whether temporal correlations may ultimately help operators detect the underlying causes for performance fluctuations or outright failure. For example, correlated failures could ultimately allow operators to trace failures to a common DSLAM, transit link, etc. Second, our findings might ultimately improve reliability for access networks.

REFERENCES

[1] A. Akella, B. Maggs, S. Seshan, A. Shaikh, and R. Sitaraman. A measurement-based analysis of multihoming. In *Proc. ACM SIGCOMM*, Karlsruhe, Germany, Aug. 2003.

[2] A. Akella, J. Pang, B. Maggs, S. Seshan, and A. Shaikh. A comparison of overlay routing and multihoming route control. In *Proc. ACM SIGCOMM*, Portland, OR, Aug. 2004.

[3] D. G. Andersen, H. Balakrishnan, M. F. Kaashoek, and R. Rao. Improving Web availability for clients with MONET. In *Proc. 2nd USENIX NSDI*, Boston, MA, May 2005.

[4] J. Bauer. Unbundling Policy in the United States: Players, Outcomes and Effects. Technical Report 05-02, Michigan State University, Department of Telecommunication, Mar. 2005. `http://www.ictregulationtoolkit.org/en/Document.2904.pdf`.

[5] R. Carlson. Network Diagnostic Tool. `http://e2epi.internet2.edu/ndt/`.

[6] K. Cho, K. Fukuda, H. Esaki, and A. Kato. The impact and implications of the growth in residential user-to-user traffic. In *ACM SIGCOMM 2006*, 2006.

[7] D. Croce, T. En-Najjary, G. Urvoy-Keller, and E. Biersack. Capacity estimation of adsl links. In *CoNEXT*, 2008.

[8] Dégroupage Free: état des DSLAM er connexions ADSL. `http://www.francois04.free.fr/liste_dslam.php`.

[9] M. Dischinger, A. Haeberlen, K. P. Gummadi, and S. Saroiu. Characterizing residential broadband networks. In *Proc. ACM SIGCOMM Internet Measurement Conference*, San Diego, CA, USA, Oct. 2007.

[10] M. M. et al. Network Path and Application Diagnosis. `http://www.psc.edu/networking/projects/pathdiag/`.

[11] J. Friedman and B. Popescu. Gradient directed regularization. *Stanford University, Technical Report*, 2003.

[12] J. Friedman and B. Popescu. Predictive learning via rule ensembles. *Annals of Applied Statistics (to appear)*, 2008.

[13] Glasnost: Bringing Transparency to the Internet. `http://broadband.mpi-sws.mpg.de/transparency`.

[14] Grenouille. Grenouille. `http://www.grenouille.com/`.

[15] D. Han, A. Agarwala, D. G. Andersen, M. Kaminsky, K. Papagiannaki, and S. Seshan. Mark-and-sweep: Getting the inside scoop on neighborhood networks. In *Proc. Internet Measurement Conference*, Vouliagmeni, Greece, Oct. 2008.

[16] Internet World Stats. `http://www.internetworldstats.com/dsl.htm`.

[17] Asymmetric Digital Subscriber Line (ADSL) Transceivers. ITU-T G.992.1, 1999.

[18] Asymmetric Digital Subscriber Line (ADSL) Transceivers - Extended Bandwidth ADSL2 (ADSL2Plus). ITU-T G.992.5, 2003.

[19] C. R. S. Jr. and G. F. Riley. Neti@home: A distributed approach to collecting end-to-end network performance measurements. In *the Passive and Active Measurement Conference (PAM)*, 2004.

[20] Keynote Home Page. `http://www.keynote.com/`, 1999.

[21] C. Kuang. Burning Question: Does Internet Speed Vary by Season? `http://www.wired.com/gadgets/miscellaneous/magazine/17-10/ts_burningquestion`, Sept. 2009.

[22] L'autorité de Régulation des Communications Életroniques et des Postes. Observatoires/Tableau de bord dégroupage et bitstream. `http://www.acerp.fr/index.php?id=9568`.

[23] R. Mahajan, M. Zhang, L. Poole, and V. Pai. Uncovering Performance Differences among Backbone ISPs with Netdiff. In *Proc. 5th USENIX NSDI*, San Francisco, CA, Apr. 2008.

[24] G. Maier, A. Feldmann, V. Paxson, and M. Allman. On dominant characteristics of residential broadband internet traffic. In *ACM Internet Measurement Conference*, 2009.

[25] Netalyzr. `http://netalyzr.icsi.berkeley.edu/`.

[26] M. Siekkinen, D. Collange, G. Urvoy-Keller, and E. Biersack. Performance limitations of adsl users: A case study. In *the Passive and Active Measurement Conference (PAM)*, 2007.

[27] M. B. Tariq, M. Motiwala, and N. Feamster. Detecting Network Neutrality Violations with Causal Inference. In *Proc. CoNEXT*, Dec. 2009.